TITLE OF THE INVENTION

DATA TRANSMISSION APPARATUS AND METHOD

BACKGROUND OF THE INVENTION

Field of the Invention

[0001] The present invention relates to apparatuses and methods for data transmission on a packet basis and, more specifically, to a data transmission apparatus and method wherein every packet is assigned with a priority, and only any packet high in priority is to be retransmitted.

Description of the Background Art

[0002] Data such as video and audio is generally transmitted in real-time under a Realtime Transport Protocol (RTP). The details of the RTP are found in "RTP: A Transport Protocol for Real-Time Applications", H. Schulzrinne, S. Casner, R. Frederik, and V. Jacobson, RFC1889, 1996.

[0003] FIG. 9 is a block diagram showing the configuration of a data transmission apparatus having the RTP applied. In FIG. 9, a data transmitter 50 and a data receiver 60 work together to transmit data from a transmission-side application 1 to a reception-side application 2. Herein, the transmission-side application 1 is exemplified by a video encoder and audio encoder, while the reception-side application 2 by a corresponding video decoder and audio decoder.

[0004] A header adding part 51 provides a packet outputted from the transmission-side application 1 with a header including a sequence number and a timestamp, for example. A packet transmitter 52 transmits a data packet 101 with such header added. A packet receiver 62 receives and outputs the data packet 101 to a packet output part 61, wherein the data packet is stored. The packet output part 61 outputs the stored data packet to the reception-side application 2 based on the timestamp.

[0005] During data transmission over a general transmission path, packet loss is likely to occur due to transmission error and congestion. Under the RTP, the packet receiver 62 refers to the sequence numbers assigned to each of the received data packets 101 to detect if any packet loss has been occurred. How the packet receiver 62 is receiving the data packets (hereinafter, packet reception state) is informed to a reception state transmitter 63. The reception state transmitter 63 then generates a receiver report packet (RR packet) 110, indicating the largest sequence number found among those so far received, and the total number of packets lost so far. Thus generated RR packet 110 is transmitted to a reception state receiver 53, and the reception state found in the RR packet 110 is then notified to both the transmission-side application 1 and the packet transmitter 52.

[0006] FIG. 10 is a sequence diagram for data transmission under the RTP. The data transmitter 50 assigns a sequence number

(SN) to every packet (denoted by an arrow), and transmits the packets sequentially. The data receiver 60 informs the data transmitter 50 of the packet reception state by transmitting an RR packet with a predetermined timing.

[0007] Also known is data transmission under an enhanced RTP (priority-assigned RTP) wherein every packet is assigned with a priority, and only any packet high in priority is to be retransmitted. FIG. 11 is a block diagram showing the configuration of a data transmission apparatus having such priority-assigned RTP applied. A priority assigning part 72 sequentially assigns each header-added packet with a priority of high or low, and a second sequence number. Here, the second sequence number is incremented by 1 whenever the subsequent packet is of high priority, but remains the same if the packet low in priority follows. A packet transmitter 73 transmits a data packet 100 with such priority assigned.

[0008] A priority judging part 82 refers to the second sequence numbers to detect whether any packet of high priority has been lost. The detection result is informed to a transmission request transmitter 85. In response, the retransmission request transmitter 85 transmits to a transmission request receiver 75 a negative acknowledge packet (NACK packet) 120 designating which packet needs to be retransmitted. The retransmission request receiver 75 then passes the information to a retransmitting packet supply part 76, wherein the packets of high priority are stored

for retransmission. Based on the information, the retransmitting packet supply part 76 responsively outputs any designated packet to the priority assigning part 72. As such, under the priority-assigned RTP, only any packet of high priority is to be processed for retransmission.

[0009] FIG. 12 is a sequence diagram for data transmission under the priority-assigned RTP. In FIG. 12, reference characters SN and SSN denote, respectively, the sequence number and the second sequence number. Moreover, "P=1" denotes a packet of high priority, while "P=0" a packet low in priority. In FIG. 12, for example, when packets of SN=3 and SN=4 are both lost, the latter packet is to be retransmitted as is high in priority, but not the former packet of low priority.

[0010] Here, in the above, priority assignment is performed with no consideration of the packet reception state, and a ratio between high and low priorities to be assigned is constant. Therefore, no matter if the transmission path is bad in condition and packet loss often occurs, packet retransmission continues, resulting in increase of data for transmission. On the other hand, even if the transmission capacity is ample with the transmission path being good in condition and thus rare packet loss, the packets of high priority are generated yet based on the constant ratio, and packet retransmission is performed only when any packet of high priority is lost. As such, due to the neglect of the packet reception state at the time of priority assignment, data

transmission cannot be efficiently carried out with whatever available transmission capacity.

SUMMARY OF THE INVENTION

[0011] Therefore, an object of the present invention is to provide a data transmission apparatus and method achieving efficient data transmission with the appropriate manner change of priority assignment in consideration of the packet reception state.

[0012] The present invention has the following features to attain the object above.

[0013] A first aspect of the present invention is directed to a data transmission apparatus for transmitting data on a packet basis from a transmitter to a receiver. The transmitter comprises: a priority assigning part for assigning priorities on a packet basis; a packet transmission part for transmitting a priority-assigned packet; a reception state receiving part for receiving a packet reception state in the receiver; and a packet retransmitting part for performing packet retransmission in response to a retransmission request from the receiver. The receiver comprises: a packet reception part for receiving the packet provided by the packet transmission part; a reception state transmitting part for transmitting the reception state based on packet loss information detected by the packet reception part; and a retransmission request transmitting part for transmitting

the retransmission request if detected any packet of high priority as having been lost. Herein, the priority assignment part changes manners of priority assignment so that the packet of higher priority is decreased in number when the reception state is bad, and increased in number when good.

[0014] As described above, in the first aspect, any packet of high priority is generated less when the packet reception state in the receiver is bad, and generated more when the packet reception state is good. As such, by controlling the amount of data to be retransmitted based on the reception state, data transmission can be efficiently carried out with whatever available transmission capacity.

[0015] More preferably, the reception state transmission part may transmit the reception state including a packet loss ratio in the packet reception part, and the priority assignment part may change the manners of priority assignment so that the packet of high priority is found with a lower ratio when the packet loss ratio is larger than a predetermined value, and when smaller, the packet of high priority is found with a higher ratio.

[0016] According to a second aspect, in the first aspect, the priority assigning part classifies any packet including coded data derived from moving pictures into an intra-coded packet carrying intra-frame coded data, or an inter-coded packet carrying inter-frame coded data, and based on the reception state, changes the manners of priority assignment determined according

to packet type.

[0017] As described above, in the second aspect, to transmit coded data derived from moving pictures, the packets are classified into two types in consideration of any influence to be exerted upon the resulting reproduced image. Based on the classification and the packet reception state in the receiver, packets are assigned with priorities. As such, by controlling the amount of data to be retransmitted based on the reception state, data transmission can be efficiently carried out with whatever available transmission capacity, preventing quality deterioration to any resulting reproduced image.

[0018] In such case, the priority assigning part may change between, based on the reception state, a first priority assigning manner wherein the intra-coded packet is assigned with a high priority, and the inter-coded packet with a low priority, and a second priority assigning manner wherein every packet is assigned with a high priority. Or alternatively, based on the reception state, changed among may be a first priority assigning manner wherein the intra-coded packet is assigned with either a high or low priority at a predetermined ratio and the inter-coded packet with a low priority, a second priority assigning manner wherein the intra-coded packet is assigned with a high priority and the inter-coded packet with a low priority, and a third priority assigning manner wherein every packet is assigned with a high priority.

[0019] When coded data derived from moving pictures is transmitted, the packets are classified into two types in consideration of any effect to be exerted upon the resulting reproduced image. In this manner, by changing two or three manners of priority assignment, it becomes possible to prevent quality deterioration to any resulting reproduced image.

[0020] A third aspect of the present invention is directed to a data transmission method for transmitting data on a packet basis from a transmitter to a receiver. The transmitter comprises the steps of: assigning priorities on a packet basis; transmitting a priority-assigned packet; receiving a packet reception state in the receiver; and performing packet retransmission in response to a retransmission request from the receiver. The receiver comprises the steps of: receiving the packet provided by the packet transmission part; transmitting the reception state based on packet loss information detected in the packet receiving step; and transmitting the retransmission request if detected any packet of high priority as having been lost. Herein, in the priority assigning step, manners of priority assignment is so changed that the packet of higher priority is decreased in number when the reception state is bad, and increased in number when good.

[0021] As described above, in the third aspect, any packet of high priority is generated less when the packet reception state in the receiver is bad, and generated more when the packet reception state is good. As such, by controlling the amount of

data to be retransmitted based on the reception state, data transmission can be efficiently carried out with whatever available transmission capacity.

[0022] More preferably, in the reception state transmitting step, the reception state including a packet loss ratio in the packet receiving step may be transmitted, and in the priority assigning step, the manners of priority assignment may be so changed that the packet of high priority is found with a lower ratio when the packet loss ratio is larger than a predetermined value, and when smaller, the packet of high priority is found with a higher ratio.

[0023] According to a fourth aspect, in the third aspect, in the priority assigning step, any packet including coded data derived from moving pictures is classified into an intra-coded packet carrying intra-frame coded data or an inter-coded packet carrying inter-frame coded data, and based on the reception state, the manners of priority assignment determined according to packet type are changed.

[0024] As described above, in the fourth aspect, to transmit coded data derived from moving pictures, the packets are classified into two types in consideration of any influence to be exerted upon the resulting reproduced image. Based on the classification and the packet reception state in the receiver, packets are assigned with priorities. As such, by controlling the amount of data to be retransmitted based on the reception state,

data transmission can be efficiently carried out with whatever available transmission capacity, preventing quality deterioration to any resulting reproduced image.

[0025] In this case, in the priority assigning step, based on the reception state, changed between may be a first priority assigning manner wherein the intra-coded packet is assigned with a high priority and the inter-coded packet with a low priority, and a second priority assigning manner wherein every packet is assigned with a high priority. Or alternatively, based on the reception state, changed among may be a first priority assigning manner wherein the intra-coded packet is assigned with either a high or low priority at a predetermined ratio and the inter-coded packet with a low priority, a second priority assigning manner wherein the intra-coded packet is assigned with a high priority and the inter-coded packet with a low priority, and a third priority assigning manner wherein every packet is assigned with a high priority.

[0026] When coded data derived from moving pictures is transmitted, the packets are classified into two types in consideration of any effect to be exerted upon the resulting reproduced image. In this manner, by changing two or three manners of priority assignment, it becomes possible to prevent quality deterioration to any resulting reproduced image.

[0027] These and other objects, features, aspects and advantages of the present invention will become more apparent from

the following detailed description of the present invention when taken in conjunction with the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

[0028] FIG. 1 is a block diagram showing the configuration of a data transmission apparatus according to an embodiment of the present invention;

FIG. 2 is a diagram showing the format of a data packet under a priority-assigned RTP;

FIG. 3 is a table showing manners of priority assignment differed among priority-assigning parts in the data transmission apparatus of the embodiment;

FIG. 4 is a diagram showing the format of a NACK packet under the priority-assigned RTP;

FIG. 5 is a diagram showing the format of an RR packet under the priority-assigned RTP;

FIG. 6 is a diagram showing the manner change of priority assignment in the data transmission apparatus of the embodiment;

FIG. 7 is a block diagram exemplarily showing the modified configuration of the data transmission apparatus according to the embodiment of the present invention;

FIG. 8 is a table showing manners of priority assignment differed between priority-assigning parts in the data transmission apparatus of the embodiment;

FIG. 9 is a block diagram showing the configuration of

a conventional data transmission apparatus having an RTP applied;

FIG. 10 is a sequence diagram for data transmission under the RTP;

FIG. 11 is a block diagram showing the configuration of the conventional data transmission apparatus under the priority-assigned RTP; and

FIG. 12 is a sequence diagram for data transmission under the priority-assigned RTP.

DESCRIPTION OF THE PREFERRED EMBODIMENT

[0029] FIG. 1 is a block diagram showing the configuration of a data transmission apparatus according to an embodiment of the present invention. This data transmission apparatus is configured by a data transmitter 10 and a data receiver 20, and data is transmitted from the transmission-side application 1 to the reception-side application 2 under the priority-assigned RTP. The data transmitter 10 includes a header adding part 11, a priority assignment control part 12, first to third priority assigning parts 13a to 13c, a packet transmitter 14, a reception state receiver 15, a retransmission request receiver 16, and a retransmitting packet supply part 17. As to the data receiver 20, included are a packet output part 21, a priority judging part 22, a packet receiver 23, a packet reception state transmitter 24, and a retransmission request transmitter 25. Herein, any component appeared in the apparatus of FIG. 9 is under the same

reference numeral.

[0030] Described next is data transmission from the transmission-side application 1 to the reception-side application 2. First, the header adding part 11 provides a packet outputted from the transmission-side application 1 with a header including a sequence number and a timestamp, for example. The priority assignment control part 12 outputs the header-added packet to any of the first to third priority assigning part 13a to 13c. As will be later described, the first to third priority assigning parts 13a to 13c provide every incoming packet with a priority of high or low in each different manner. The packet transmitter 14 transmits the priority-assigned data packet 100 to the packet receiver 23.

[0031] The data packet 100 thus received by the packet receiver 23 is then forwarded to the packet output part 21 via the priority judging part 22. The packet output part 21 stores the data packet 100, and waits for the time designated by its timestamp for output to the reception-side application 2. The priority judging part 22 detects whether any packet of high priority has been lost, and the detection result is notified to the retransmission request transmitter 25. The retransmission request transmitter 25 transmits the NACK packet 120 designating which packet is to be retransmitted.

[0032] Based on the received NACK packet 120, the retransmission request receiver 16 informs the retransmitting

packet supply part 17 which packet needs to be retransmitted. Here, the retransmitting packet supply part 17 previously stores any packet assigned with a high priority by the first to third priority assigning parts 13a to 13c, and responsively outputs the designated packet to the priority assignment control part 12.

[0033] The packet receiver 23 also outputs the data packet 100 to the reception state transmitter 24 after detecting any packet loss. Based on the information about packet loss, the reception state transmitter 24 generates the RR packet 110, and transmits it to the reception state receiver 15. From the RR packet 110, the reception state receiver 15 extracts a packet loss ratio 200, and outputs it to the priority assignment control part 12. Based on the packet loss ratio 200, the priority assignment control part 12 determines to which of the first to third priority assigning parts 13a to 13c the packets each provided by the header adding part 11 and the retransmitting packet supplying part 17 are to be outputted.

[0034] Next, the data transmission apparatus of the present embodiment is described in detail.

[0035] Here, the transmission-side application 1 is presumably an application for encoding moving pictures such as MPEG 4 and H. 263, for example. The transmission-side application 1 first performs intra-coding or inter-coding with respect to moving pictures, and the resultingly derived coded data is divided into packets. The packets are classified into

"intra-coded packets" carrying intra-frame coded data, and "inter-coded packets" carrying inter-frame coded data. In comparison, the intra-coded packets exert more significant influence upon the quality of the resulting reproduced image than the inter-coded packets, and thus considered more consequential. Such packet characteristic is outputted from the transmission-side application 1 together with the packets.

[0036] In the present embodiment, presumably, the intra-coded packet corresponds to any intra-coded frame, and the inter-coded packet to any inter-coded frame. Alternatively, any packet carrying the predetermined number or more of intra-coded blocks may be handled as the intra-coded packet, or any packet carrying the intra-coded blocks with a predetermined ratio or more may be handled as such. In this case, any other packet may be handled as the inter-coded packet.

[0037] FIG. 2 is a diagram showing the format of the priority-assigned data packet 100. In FIG. 2, a payload PL is packet data outputted from the transmission-side application 1. A payload type PLT denotes the priority-assigned RTP as being applied for data transmission. A sequence number SN is a number sequentially incremented by 1 in order of data packets. A timestamp TS denotes a time when the packet is to be used in the reception-side application 2. A transmission-side identifier SSRC identifies the data transmitter 10. A priority P denotes the packet priority, and a value 1 means the

packet as being high in priority, and a value 0 as being low in priority. A second payload type PLT2 denotes the type of data to be transmitted under the priority-assigned RTP. A second sequence number SSN is a number incremented by 1 whenever the subsequent packet is of high priority, but remains the same for the packet low in priority.

[0038] The header adding part 11 provides the packet outputted from the transmission-side application 1 with a header, and sets a value to fields from the head of the header to the transmission-side identifier. The header-added packet is outputted by the priority assignment control part 12 to any of the first to third priority assigning parts 13a to 13c.

[0039] As shown in FIG. 3, the first to third priority assigning parts 13a to 13c perform priority assignment in each different manner. In detail, in the first priority assigning part 13a, the priority assigned to the intra-coded packets is either high or low with a ratio of 1:1, and the priority to the inter-coded packets is low. In the second priority assigning part 13b, the priority assigned to the intra-coded packets is high, but to the inter-coded packets is low. The third priority-assigning part 13c assigns every incoming packet with a low priority. Based on the priorities assigned as such, the first to third priority assigning parts 13a to 13c assign each packet with the second sequence number.

[0040] The packet transmitter 14 transmits the priority-

assigned data packets 100 provided by the first to third priority assigning parts 13a to 13c. The data packets 100 are transmitted over the Internet using the User Datagram Protocol (UDP) module (not shown), for example.

[0041] In response to the data packets 100, the packet output part 21 and the packet receiver 23 operate as described above. The priority judging part 22 refers to the priorities and the second sequence numbers to detect whether any packet of high priority has been lost, and notifies the detection result to the retransmission request transmitter 25. Specifically, if detecting any packet of high priority having the second sequence number larger by 2 or more than that of the preceding packet, or if detecting any packet of low priority having the second sequence number different from that of the preceding packet, the priority judging part 22 determines that any packet of high priority has been lost.

[0042] The retransmission request transmitter 25 generates, based on the notification from the priority judging part 22, the NACK packet 120 designating which packet needs to be retransmitted. FIG. 4 is a diagram showing the format of the NACK packet. In FIG. 4, a packet type (PT) denotes the type of packet. A packet length (L) denotes the length of the packet. A reception-side identifier (SSRC_R) is an identifier for identifying the data receiver 20. The transmission-side identifier SSRC identifies from where the data packet 100 came, that is, the data

transmitter 10. The second sequence number (SSN) is the second sequence number of the packet to be retransmitted. Here, the NACK packet shown in FIG. 4 is designating only one packet for retransmission. The number of the packets to be designated is not restrictive, and several packets may be simultaneously designated by including any appropriate field for other data.

[0043] The retransmission request transmitter 25 transmits thus generated NACK packet 120 over the Internet using the UDP module (not shown), for example, as is the data packet 100.

[0044] In response to the NACK packet 120, the retransmission request receiver 16 and the retransmitting packet supply part 17 operate as described above. The first to third priority assigning parts 13a to 13c, and the packet transmitter 14 execute the processing to the packet provided by the retransmitting packet supply part 17 in the same manner as to the packet coming from the header adding part 11. Accordingly, any packet of high priority designated by the NACK packet 120 is retransmitted.

[0045] The packet receiver 23 detects any packet loss by referring to the sequence numbers of the received data packets 100. If detecting any packet of high priority having the sequence number larger by 2 or more than that of the preceding packet, the packet receiver 23 determines as packet loss having been occurred. The packet receiver 23 notifies information about packet reception state and packet loss to the reception state transmitter 24.

[0046] Based on the notified information, the reception state transmitter 24 calculates the total number of packets lost so far and the packet loss ratio, and generates the RR packet 110 including thus calculated values. FIG. 5 is a diagram showing the format of the RR packet 110. Therein, the packet type (PT), the packet length (L), the reception-side identifier (SSRC_R), and the transmission-side identifier (SSRC) are the same as those in the format of the NACK packet 120. Here, a cumulative number of packets lost (CNPL) and a fraction lost (FL: equal to packet loss ratio) indicate the values calculated by the reception state transmitter 24. An extended highest sequence number received (EHSNR) denotes the sequence number of the packet last received. An interarrival jitter (IJ) denotes a jitter observed for the delay over the transmission path, and a last SR (LSR) and a delay since last SR (DLSR) are used to measure the round-trip delay over the transmission path. Here, the details of each field are also found in the above document (RFC1889).

[0047] The reception state transmitter 24 transmits thus generated RR packet 110 over the Internet by using the UDP module (not shown), for example, as is the data packet 100. The reception state receiver 15 extracts the packet loss ratio 200 from the RR packet 110, and outputs it to the priority assignment control part 12.

[0048] Based on the packet loss ratio 200, the priority assignment control part 12 changes to which of the first to third

priority assigning parts 13a to 13c the packets provided by the header adding part 11 and the retransmitting packet supply part 17 are to be outputted. Here, assuming that the packet loss ratio 200 is X(%), and threshold values for the change are T1 and T2. When X is T1 or larger, the priority assignment control part 12 outputs the received packet to the first priority assigning part 13a. When X is T2 or lager and smaller than T1, the received packet goes to the second priority assigning part 13b, and when X is smaller than T2, to the third priority assigning part 13c. As an example, when T1 is 70% and T2 is 30%, the priority assignment control part 12 outputs the received packet to the first priority assigning part 13a with X being 70% or more, to the second priority assigning part 13b with X being 30% or larger and smaller than 70%, and to the third priority assigning part 13c with X being smaller than 30%. As such, the manners of priority assignment are changed based on the packet loss ratio.

[0049] FIG. 6 is a diagram showing the manner change of priority assignment done by the data transmitter 10 after receiving the RR packet 110. For the sake of simplification, FIG. 6 shows only the RR packet 110. The data receiver 20 transmits, with a predetermined timing, the RR packet 110 carrying the packet loss ratio. In response to the packet loss ratio, the data transmitter 10 changes the manners of priority assignment. Specifically, the manners are so changed that the number of packets high in priority becomes less with the high packet loss

ratio, and with the low packet loss ratio, the number of packets high in priority are increased. Accordingly, packet retransmission is suppressed with the high packet loss ratio, whereby the data to be transmitted is reduced. On the other hand, packet retransmission is encouraged with the low packet loss ratio, whereby the data to be transmitted is increased. As such, by changing the manners of priority assignment and controlling data amount for retransmission, coded data derived from moving pictures can be efficiently transmitted with whatever available data capacity, preventing quality deterioration to any resulting reproduced image.

[0050] As such, in the data transmission apparatus of the present embodiment, the manners of priority assignment are changed on the transmission side in consideration of the packet reception state, and thus the packets of high quality are increased in number when the packet reception state is bad, and decreased in number with good packet reception state. In this manner, the data to be retransmitted can be controlled in amount, and thus data transmission can be efficiently performed with whatever available data capacity.

[0051] Here, the first to third priority assigning parts 13a to 13c are provided in the data transmission apparatus, and priority assignment is changed in three different manners. This is not restrictive, and as shown in FIG. 7, for example, the data transmission apparatus may include first and second priority

assigning parts 33a and 33b, and these may operate according to a table shown in FIG. 8.

[0052] Further, the packet characteristic is outputted from the transmission-side application in the above. However, the data transmitter may calculate that. For example, to transmit image coded data under the RTP, the packet characteristic may be easily derived by referring to a specific bit in the packet.

[0053] Still further, data to be transmitted is arbitrary, and so is the manner of packet classification. For example, to transmit data derived by hierarchically coding the image in the frequency direction, the same manner as above can be applied by classifying any packet including a low frequency component into the intra-coded packet, or any packet including a high frequency component into the inter-coded packet. Also for transmission of audio data, any packet carrying real sound may be classified into the intra-coded packet, or any packet carrying silent sound for most of the part into the inter-coded packet. As another possibility, priority assignment may be done without referring to the packet characteristic.

[0054] In the present embodiment, the communications protocol is the RTP. This is not restrictive, and any other communications protocol will also do, leading to the data transmission apparatus with the same features as above.

[0055] While the invention has been described in detail, the foregoing description is in all aspects illustrative and not

restrictive. It is understood that numerous other modifications and variations can be devised without departing from the scope of the invention.